BIOLOGICAL CRYSTALLOGRAPHY

ISSN 1399-0047

Against the odds? *De novo* structure determination of a pilin with two cysteine residues by sulfur SAD

Manuela Gorgel,[a] Andreas Bøggild,[a] Jakob Jensen Ulstrup,[a] Manfred S. Weiss,[b] Uwe Müller,[b] Poul Nissen[a] and Thomas Boesen[a]*

[a]Department of Molecular Biology and Genetics, Aarhus University, Gustav Wieds Vej 10C, DK-8000 Aarhus C, Denmark, and [b]Macromolecular Crystallography (HZB-MX), Helmholtz Zentrum Berlin für Materialien und Energie, Albert-Einstein-Strasse 15, D-12489 Berlin, Germany. *Correspondence e-mail: thb@mbg.au.dk

Received 15 January 2015
Accepted 16 February 2015

Edited by M. Schiltz, Fonds National de la Recherche, Luxembourg

Keywords: pilin; S-SAD; *de novo* structure determination.

PDB reference: PilBac1, 4us7

Supporting information: this article has supporting information at journals.iucr.org/d

Exploiting the anomalous signal of the intrinsic S atoms to phase a protein structure is advantageous, as ideally only a single well diffracting native crystal is required. However, sulfur is a weak anomalous scatterer at the typical wavelengths used for X-ray diffraction experiments, and therefore sulfur SAD data sets need to be recorded with a high multiplicity. In this study, the structure of a small pilin protein was determined by sulfur SAD despite several obstacles such as a low anomalous signal (a theoretical Bijvoet ratio of 0.9% at a wavelength of 1.8 Å), radiation damage-induced reduction of the cysteines and a multiplicity of only 5.5. The anomalous signal was improved by merging three data sets from different volumes of a single crystal, yielding a multiplicity of 17.5, and a sodium ion was added to the substructure of anomalous scatterers. In general, all data sets were balanced around the threshold values for a successful phasing strategy. In addition, a collection of statistics on structures from the PDB that were solved by sulfur SAD are presented and compared with the data. Looking at the quality indicator $R_{\rm anom}/R_{\rm p.i.m.}$, an inconsistency in the documentation of the anomalous R factor is noted and reported.

1. Introduction

Single-wavelength anomalous diffraction (SAD) phasing is based on differences in intensities between Friedel mates recorded at a single wavelength (Dauter *et al.*, 2002). Sulfur and phosphorus are the lightest atoms occurring in biomolecules for which anomalous signals can be used for SAD phasing without any need for derivatization or a homologous structure. However, the anomalous signal reaches only approximately 1% of the overall intensities, even at the wavelengths of 1.5–2.3 Å available at several sources. Therefore, sulfur and phosphorus SAD data sets have to be measured very precisely, which is typically obtained by high multiplicity. Owing to these limitations, only a low number of sulfur SAD (S-SAD) structures have been deposited in the PDB (104 out of 105 499; Supplementary Table S1).

Recently, highly redundant data sets (up to 535) for S-SAD have been obtained using multiple crystals and have allowed the determination of structures which could not be achieved based on single-crystal data sets (Liu, Dahmane *et al.*, 2012; Liu *et al.*, 2014). This highlights the importance of accuracy in SAD on multiple crystals. However, this approach necessitates isomorphism between crystals and is therefore not always possible.

In this study, we phased the crystal structure of a pilin dimer from *Shewanella oneidensis* ($\rm Pil_{Bac1}$) despite severe radiation

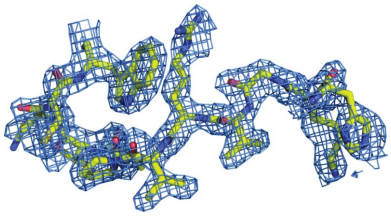


Table 1
Data-collection and processing statistics.

Values in parentheses are for the highest resolution shell. For data processing, $I+$ and $I-$ were kept separate. R_{merge} and R_{meas} were taken from *XDS* and $R_{p.i.m.}$ was calculated by dividing R_{meas} by the square root of the multiplicity.

	A1	A2	A3	A1–3	B1	B2	B3	B1–3
Beamline	I24, Diamond				14.1, BESSY II			
Detector	Pilatus 6M				Pilatus 6M			
Wavelength (Å)	1.77124				1.8			
Exposure time (s)	0.2	0.2	0.2		0.7	0.5	0.5	
Crystal-to-detector distance (mm)	280.0				149.2			
Oscillation range (°)	0.5				0.1			
Resolution	48.17–2.70 (2.85–2.70)	48.17–2.70 (2.85–2.70)	48.19–2.70 (2.84–2.70)	48.19–2.70 (2.84–2.70)	49.28–1.96 (2.01–1.96)	48.26–1.96 (2.01–1.96)	48.22–1.96 (2.01–1.96)	48.27–1.96 (2.03–1.96)
Space group	*I*222							
Unit-cell parameters								
a (Å)	48.74	48.81	48.82	48.74	48.83	48.83	48.84	48.83
b (Å)	96.35	96.34	96.37	96.35	96.55	96.51	96.44	96.55
c (Å)	109.71	109.89	109.97	109.71	109.76	109.76	109.76	109.76
$\alpha = \beta = \gamma$ (°)	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0
Mosaicity (°)	0.117	0.111	0.123		0.079	0.084	0.079	
No. of unique reflections	13663 (996)	13655 (987)	13701 (974)	13696 (1008)	34575 (2454)	34853 (2492)	34658 (2499)	34976 (2493)
Total No. of reflections	92398 (5603)	92111 (5456)	91852 (5389)	276176 (16577)	210237 (11740)	213211 (12125)	189810 (10854)	611933 (34633)
Multiplicity	6.8 (5.6)	6.7 (5.5)	6.7 (5.5)	20.2 (16.5)	6.1 (4.8)	6.1 (4.9)	5.5 (4.3)	17.5 (13.9)
Completeness (%)	99.5 (98.6)	99.4 (98.2)	99.2 (96.2)	99.5 (98.5)	96.0 (93.4)	96.8 (94.6)	96.3 (94.6)	97.1 (94.9)
R_{merge}	0.050 (0.148)	0.057 (0.142)	0.039 (0.092)	0.054 (0.140)	0.033 (0.081)	0.034 (0.086)	0.030 (0.074)	0.037 (0.088)
R_{meas}	0.054 (0.162)	0.062 (0.157)	0.042 (0.102)	0.055 (0.145)	0.036 (0.091)	0.037 (0.096)	0.033 (0.085)	0.038 (0.092)
$R_{p.i.m.}$	0.021 (0.069)	0.024 (0.061)	0.016 (0.029)	0.012 (0.036)	0.015 (0.041)	0.015 (0.042)	0.014 (0.041)	0.009 (0.025)
Wilson *B* factor† (Å^2)	46.8	32.7	41.5	34.5	28.6	27.8	28.0	27.8
Mean $I/\sigma(I)$	28.5 (9.4)	29.6 (12.1)	36.0 (14.0)	49.9 (19.9)	34.1 (13.5)	33.1 (12.9)	34.64 (13.2)	51.72 (21.2)
$CC_{1/2}$	0.999 (0.986)	0.999 (0.988)	0.999 (0.994)	1.000 (0.996)	0.999 (0.995)	0.999 (0.995)	0.999 (0.995)	1.000 (0.998)

† Taken from *CTRUNCATE* through *AIMLESS* in *CCP*4 (Evans, 2011; Evans & Murshudov, 2013).

damage, a low multiplicity of 5.5 and a low anomalous signal from only two disulfide bridges and a single sulfate ion in the asymmetric unit. By comparisons with previous S-SAD structures (Supplementary Table S1), we discuss different parameters indicating the signal strength and suggested cutoff values for successful sulfur SAD phasing.

2. Materials and methods

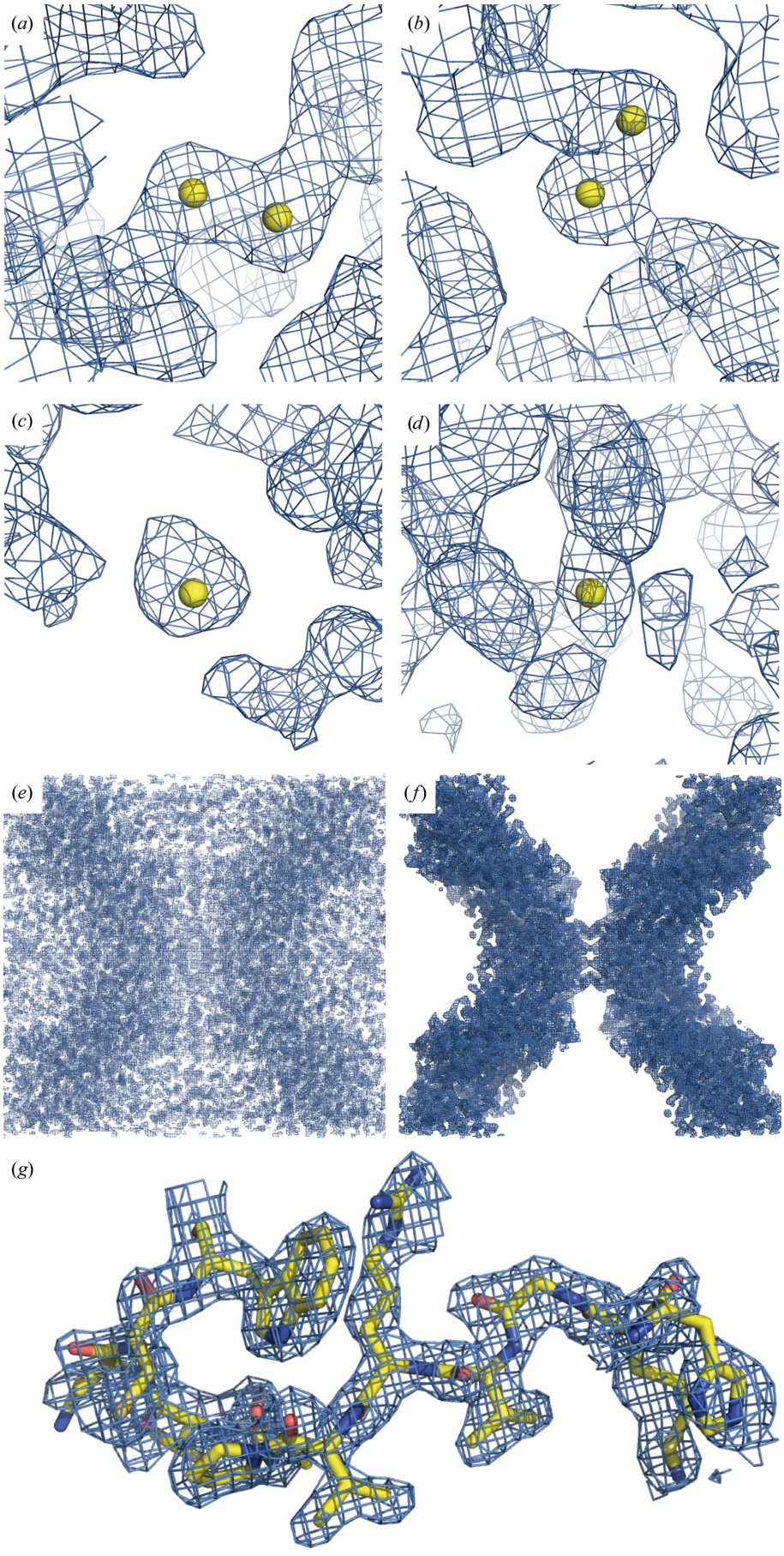
Details of the materials and methods are described in the Supporting Information. In brief, two subsets of data (A and B) were collected on two different synchrotron beamlines (Table 1). Data sets A1, A2 and A3 were collected from three different volumes of a single crystal at a wavelength of 1.77 Å on beamline I24, Diamond Light Source, England at 100 K. Data sets B1, B2 and B3 were collected at a wavelength of 1.8 Å on beamline 14.1, BESSY II, Berlin, Germany (Mueller *et al.*, 2012) in a similar way, *i.e.* by exposing three different volumes of a single crystal. All data were processed with *XDS* (Kabsch, 2010) and phasing was performed with *AutoSol* from *PHENIX* (Terwilliger *et al.*, 2009). The initial model was built by *AutoBuild* and refined by *phenix.refine* (Afonine *et al.*, 2012). Each round of refinement was followed by manual model correction in *Coot* (Emsley *et al.*, 2010). The coordinates of the final model were deposited in the PDB as entry 4us7 (http://www.rcsb.org; Berman *et al.*, 2000).

3. Results and discussion

3.1. Phasing

All single data sets were of high quality as judged by the *R* factors and $CC_{1/2}$ values. Notably, the resolution of all data sets was limited by the detector position, which is reflected by average signal-to-noise levels of around 10 in the highest resolution shells (see Table 1). Matthews coefficient analysis favoured three over two molecules in the asymmetric unit (Adams *et al.*, 2010). Therefore, we searched for six sites corresponding to three disulfide bridges with an f'' of 0.76 e$^-$. This approach yielded successful solutions for data set B3 and (depending on the random seed) B2, as well as for both of the merged data sets A1–A3 and B1–B3.

Using B1–B3, *AutoSol* in fact found six sites (Figs. 1*a*–1*d*). Sites 1 and 2 as well as 3 and 4 were completely covered with electron density and were arranged in pairs with interatomic distances of 2.1 Å, fitting well with the distance between two S atoms in a disulfide bridge. Sites 5 and 6 were individual sites that were not connected to electron-density features of the protein backbone, indicative of bound molecules such as a sulfate ion or a metal ion. The arrangement of these sites indicated a correct solution, but with only two molecules in the asymmetric unit and a solvent content of approximately 62%. However, the high solvent content was beneficial for the phase refinement by density modification (see Figs. 1*e* and 1*f*), and the density-modified map exhibited a high degree of contrast and connectivity, revealing most features of the side chains



(Fig. 1*g*), and 90% of all residues were readily assigned by *AutoBuild* (Terwilliger *et al.*, 2008). A complete structure was built with the merged data set B1–B3 and refined at 1.96 Å resolution to an R_{work} and an R_{free} of 17.9 and 22%, respectively (Supplementary Table S2). In the final model the first four sites match two disulfide bridges, while the fifth site was assigned to a sulfate ion bound to residues 1 and 3 of chain *B* and site 6 to a sodium ion (Supplementary Fig. S1).

For the other data sets that produced a successful substructure determination (A1–A3, B2 and B3), the disulfides and the sulfate ion were identified as anomalous scatterers, but not the sodium ion.

3.2. Comparison of the individual data sets

3.2.1. SigAno, $CC_{1/2}^{anom}$ and $\Delta F/F$. Merging the data sets significantly increased sigAno {$[|F(+) - F(-)|/\sigma]$ from *XDS*; Kabsch, 2010} for both subsets, especially in the low-resolution range (Supplementary Fig. S2). Notably, sigAno was higher for the individual data sets from subset B than for the merged data set A1–A3, indicating higher quality of the single data sets in subset B. $CC_{1/2}^{anom}$ hinted at the higher phasing power of the merged data set A1–A3 opposed to the individual data sets (Table 2). Nevertheless, $CC_{1/2}^{anom}$ gave no indication of differences between B1, B2 and B3, and most importantly it did not indicate that the merged data set A1–A3 led to a structure solution, whereas the structure could not be solved from B1 despite its anomalous signal extending significantly

Figure 1
(*a*–*d*) Sites found by *AutoSol* (Terwilliger *et al.*, 2009) in the density-modified map contoured at 1.5σ. (*a*) Sites 1 and 2 revealed to be a disulfide bridge. (*b*) Sites 3 and 4 revealed to be a disulfide bridge. (*c*) Site 5 revealed to be a sulfate ion. (*d*) Site 6 revealed to be a sodium ion. (*e*, *f*) Overall experimental electron-density maps for data set B1–B3 before (*e*) and after (*f*) density modification. The maps were contoured at a level of 1.5σ. (*g*) Example of the quality of the density-modified map (*a*) contoured at 1.5σ and comprising residues 76–87 (HKGVRVPATCNW).

Table 2
Statistics related to the strength of the anomalous signal in the data sets.

The Bijvoet ratio was calculated using (3) (Hendrickson & Teeter, 1981; Wang *et al.*, 2006). $R_{\rm p.i.m.}$ was determined by dividing $R_{\rm r.i.m.}$ (taken as $R_{\rm meas}$ from *XDS*; Kabsch, 2010) by the square root of the multiplicity.

Phasing statistics	A1	A2	A3	A1–3	B1	B2	B3	B1–3
Resolution to which anomalous signal extends according to a significant $CC_{1/2}^{\rm anom}$† (Å)	5.4	5.4	4.9	3.8	2.6	2.6	2.6	2.5
SigAno†	0.884	0.893	0.927	1.048	0.840	0.888	0.914	1.035
No. of sites found‡	n.a.	n.a.	n.a.	5	n.a.	4	5	6
Expected $\Delta F/F$ (%)	0.92	0.92	0.92	0.92	0.94	0.94	0.94	0.94
Measured $\Delta F/F$ without centric reflections (%)	2.6	2.3	1.9	1.5	1.6	1.7	1.7	1.3
Mean dano/mean F (%)	2.2	1.9	1.6	1.3	1.4	1.5	1.5	1.1
$R_{\rm anom}$§	0.032	0.032	0.026	0.021	0.021	0.024	0.023	0.018
$R_{\rm anom}/R_{\rm p.i.m.}$¶	1.5	1.3	1.6	1.8	1.4	1.6	1.6	2.0

† Taken from *XDS* (Kabsch, 2010). ‡ Taken from *AutoSol* (Terwilliger *et al.*, 2009). § $R_{\rm anom} = \sum_{hkl} |I(hkl)| - |I(\overline{hkl})| / \sum_{hkl} \langle I(hkl) \rangle$ (Mueller-Dieckmann *et al.*, 2005). ¶ $R_{\rm p.i.m.} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$ (Weiss, 2001).

further based on a cutoff from *XDS* (Kabsch, 2010). The calculated Bijvoet ratio for both subsets is 0.9% (see below), but the observed ratios were much higher, ranging from 2.6 to 1.3% (Table 2). As noted by Dauter *et al.* (2002), this is not owing to high anomalous signal, but rather to increased noise in the data, and indeed our merged data set B1–B3 has the highest accuracy although it displays the lowest measured Bijvoet ratio.

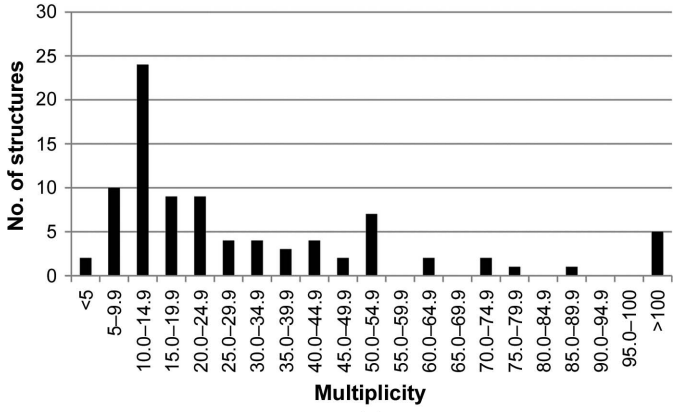
3.2.2. Anomalous difference maps. The peak heights in anomalous difference maps based on final model phases correlated well with the quality of the initial experimental maps: the higher the peaks, the better the maps (Supplementary Table S3). Despite low multiplicity, data set B3 still displayed anomalous peak heights of 22σ and 19σ, which were higher than those for the merged data sets A1–A3 (19σ and 15σ). Only for data set B did we observe the sodium ion in the anomalous difference maps (and even detected it in the substructure determination despite an f'' of 0.17 e^-). Yet, density for the sodium was evident in a $2mF_o - DF_c$ map of data set A1–A3.

Additionally, we observed a small anomalous peak close to the disulfide in data set B, indicating a reduced cysteine and thus local radiation damage (Supplementary Fig. S3). This model was confirmed by a difference map between the first and second half of each data set, which revealed positive and negative density around the cysteine bridge (chain *A*, 13–15σ; chain *B*, 15–22σ). Even though this means that the anomalous scatterer was partially shifted during the data-collection process, we were still able to detect the positions of the disulfide bridges as single-bonded sulfur positions in both molecules. The degree of local radiation damage appears to be similar for all data sets and therefore is not a cause of the phasing power difference.

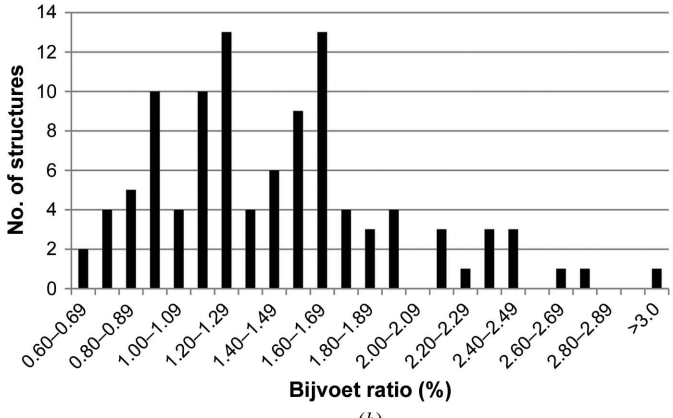
3.2.3. $R_{\rm anom}/R_{\rm p.i.m.}$. Previously, it was reported that the coefficient between the anomalous R factor $R_{\rm anom}$,

$$R_{\rm anom} = \frac{\sum_{hkl} |I(hkl)| - |I(\overline{hkl})|}{\sum_{hkl} \langle I(hkl) \rangle}, \qquad (1)$$

(Einspahr & Weiss, 2011) and the precision-indicating merging R factor $R_{\rm p.i.m.}$ should be above 1.5 to yield a consistent solution in SAD phasing. This factor was clearly above 1.5 for both merged data sets A1–A3 (1.8) and B1–B3 (2.0) and much higher than for the individual data sets, emphasizing a gain in accuracy from the merged data sets (Table 2). However, the threshold for this coefficient as an indicator of possible



(*a*)



(*b*)

Figure 2
(*a*) Numbers of structures from the PDB that were determined using the anomalous signal of sulfur, phosphorus and chloride and their multiplicities (Supplementary Table S1). (*b*) Bijvoet ratios of structures from the PDB that were solved using the anomalous signal of sulfur, phosphorus and chloride. The Bijvoet ratio was calculated as described in §3.3. We have taken the number of atoms that the structure was refined against as $N_{\rm P}$ and the number of anomalous scatterers that were deposited in the PDB as $N_{\rm A}$ (Supplementary Table S1).

structure determination is not as clear-cut in our case. The coefficients for data sets A2, A3, B1, B2 and B3 are all close to 1.5, but only data sets B2 and B3 could be solved. Therefore, we conclude that the factor $R_{\rm anom}/R_{\rm p.i.m.}$ can serve as a useful indicator, but will not necessarily identify more subtle differences in the quality of anomalous data sets.

A recently described case of S-SAD phasing was associated with an $R_{\rm anom}/R_{\rm p.i.m.}$ ratio as low as 1.1 (Huet *et al.*, 2013; Lakomek *et al.*, 2009). However, in these studies the $R_{\rm anom}$ values were determined with *SCALA* based on the definition

$$R_{\rm anom}^{SCALA} = \frac{\sum\limits_{hkl} |I(hkl)| - |I(\overline{hkl})|}{\sum\limits_{hkl} |I(hkl)| + |I(\overline{hkl})|} \qquad (2)$$

(Evans & Murshudov, 2013). In this definition of $R_{\rm anom}$ the denominator is the sum of the Friedel intensities and not the average as in (1), and $R_{\rm anom}^{SCALA}$ and subsequently $R_{\rm anom}^{SCALA}/R_{\rm p.i.m.}$ will be only about half, and the threshold value for structure determination should be 0.75. The reported cases are therefore well above the common threshold.

3.3. Comparison to S-SAD structures from the PDB

3.3.1. Multiplicity. Because of the weak anomalous signal of sulfur, data sets for S-SAD need to be recorded with high multiplicity (Dauter & Adamiak, 2001). Supplementary Table S1 shows the multiplicity for 104 S-SAD structures in the PDB, for which the average is 38 (Fig. 2). Compared with this a multiplicity of 5.5–6.8 is very low, and in fact data set B3 represents an S-SAD data set with one of the lowest multiplicities; only structures 2g51 and 2yzq have a lower multiplicity, but these proteins also display a much higher number of anomalous scatterers per protein atom (Bijvoet ratios of 1.5 and 1.6%, respectively).

3.3.2. Theroretical Bijvoet ratio. The theoretical Bijvoet ratio estimates the anomalous signal based on sequence and can be calculated as

$$\frac{\Delta F}{F} = \left(\frac{2 \sum\limits_{i} N_{{\rm A},i} f_i''^2}{N_{\rm P} Z_{\rm eff}^2} \right)^{1/2} \qquad (3)$$

(Hendrickson & Teeter, 1981), where $N_{\rm A}$ is the number of anomalous scatterers, f'' is the anomalous scattering factor, $N_{\rm P}$ is the total number of non-H atoms and $Z_{\rm eff}$ is the atomic form factor (which has an average value of 6.7 for light atoms of proteins; Hendrickson & Teeter, 1981; Wang *et al.*, 2006). The theoretical Bijvoet ratios for S-SAD structures from the PDB range from 0.6 to 4.1%, with an average of 1.5%. Based on the four cysteines and one sulfate ion, the expected Bijvoet ratios for subsets A and B at 1.77124 and 1.8 Å are 0.92 and 0.94%, respectively. Again, these values lie at the lower end of the statistics for S-SAD structures (Supplementary Table S1).

4. Conclusion

In this study, we have shown that, against the odds, the predicted limitations of sulfur SAD phasing can be challenged. Merged data sets of individual volumes of a single, well diffracting crystal can lead to powerful phasing and excellent electron-density maps even at very low Bijvoet ratios. Furthermore, we have qualified our analysis by a compilation of the sulfur SAD structures available in the PDB.

5. Related literature

The following references are cited in the Supporting Information for this article: Agarwal *et al.* (2006), Alag *et al.* (2009), Almo *et al.* (2007), Ambrosio *et al.* (2005), Araç *et al.* (2012), Bian *et al.* (2010), Bond *et al.* (2001), Brown *et al.* (2002), Chattopadhyay *et al.* (2008), Chen *et al.* (2007), Cianci *et al.* (2001), Cioci *et al.* (2011), Dauter *et al.* (1999), Debreczeni, Bunkóczi *et al.* (2003), Debreczeni, Girmann *et al.* (2003), Devesse *et al.* (2011), Doan & Dokland (2003), Enroth & Ake (2008), Fabian *et al.* (2013), Fukakusa *et al.* (2012), Gordon *et al.* (2001), Goulet *et al.* (2010), Guimarães *et al.* (2009), Ivanov *et al.* (2010), Klock *et al.* (2005), Lartigue *et al.* (2004), Li *et al.* (2002), Li *et al.* (2008), Li *et al.* (2014), Liu, Gui *et al.* (2014), Liu *et al.* (2000), Machius *et al.* (2007), Micossi *et al.* (2002), Moynie *et al.* (2013), Mueller *et al.* (2011), Mueller-Dieckmann *et al.* (2007), Müller *et al.* (2011), Ni, Benning *et al.* (2008), Ni, Sheldrick *et al.* (2008), Okada *et al.* (2008), Olsen *et al.* (2004), Pal *et al.* (2008), Phillips *et al.* (2004), Pieren *et al.* (2006), Popovic *et al.* (2012), Rajasekaran *et al.* (2013), Rodamilans *et al.* (2007), Ru *et al.* (2012), Rudiño-Piñera *et al.* (2007), Sarma & Karplus (2006), Sarma *et al.* (2010, 2012), Shinobu *et al.* (2010), Stegmann *et al.* (2009), Thomas *et al.* (2009), Thorsell *et al.* (2009), Trempe *et al.* (2005), Vasur *et al.* (2006), Watanabe *et al.* (2005), Weiss *et al.* (2004), Yogavel *et al.* (2014), Zebisch *et al.* (2013) and Zhu *et al.* (2012).

Acknowledgements

We would like to thank the Graduate School of Science and Technology, Aarhus University and the Danish National Research Foundation Centre Pumpkin – Centre for Membrane Pumps in Cells and Disease for funding the PhD project of MG. The work was supported by the advanced research program BIOMEMOS of the European Research Council (to PN) and by the Danish Council for Independent Research, Technology and Production Sciences project on microcable-based nanoelectronics (to TB). Travel costs to Diamond and BESSY were financed by Biostruct-X project 5624. We are grateful to Joseph Lyons, Aarhus University for help with data collection and George Sheldrick for discussions on the S-SAD data.

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H. & Adams, P. D. (2012). *Acta Cryst.* D**68**, 352–367.
Agarwal, R., Bonanno, J. B., Burley, S. K. & Swaminathan, S. (2006). *Acta Cryst.* D**62**, 383–391.

Alag, R., Bharatham, N., Dong, A., Hills, T., Harikishore, A., Widjaja, A. A., Shochat, S. G., Hui, R. & Yoon, H. S. (2009). *Protein Sci.* **18**, 2115–2124.

Almo, S. C. *et al.* (2007). *J. Struct. Funct. Genomics*, **8**, 121–140.

Ambrosio, A. L. B., Nonato, M. C., de Araujo, H. S. S., Arni, R., Ward, R. J., Ownby, C. L., de Souza, D. H. F. & Garratt, R. C. (2005). *J. Biol. Chem.* **280**, 7326–7335.

Araç, D., Boucard, A. A., Bolliger, M. F., Nguyen, J., Soltis, S., Südhof, T. C. & Brunger, A. T. (2012). *EMBO J.* **31**, 1364–1378.

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). *Nucleic Acids Res.* **28**, 235–242.

Bian, C., Yuan, C., Chen, L., Meehan, E. J., Jiang, L., Huang, Z., Lin, L. & Huang, M. (2010). *Proteins*, **78**, 1601–1605.

Bond, C. S., Shaw, M. P., Alphey, M. S. & Hunter, W. N. (2001). *Acta Cryst.* D**57**, 755–758.

Brown, J., Esnouf, R. M., Jones, M. A., Linnell, J., Harlos, K., Hassan, A. B. & Jones, E. Y. (2002). *EMBO J.* **21**, 1054–1062.

Chattopadhyay, K., Ramagopal, U. A., Brenowitz, M., Nathenson, S. G. & Almo, S. C. (2008). *Proc. Natl Acad. Sci. USA*, **105**, 635–640.

Chen, L., Chen, L.-R., Zhou, X. E., Wang, Y., Kahsai, M. A., Clark, A. T., Edmondson, S. P., Liu, Z.-J., Rose, J. P., Wang, B.-C., Meehan, E. J. & Shriver, J. W. (2007). *J. Mol. Biol.* **341**, 73–91.

Cianci, M., Rizkallah, P. J., Olczak, A., Raftery, J., Chayen, N. E., Zagalsky, P. F. & Helliwell, J. R. (2001). *Acta Cryst.* D**57**, 1219–1229.

Cioci, G., Terradot, L., Dian, C., Mueller-Dieckmann, C. & Leonard, G. (2011). *Proteins*, **79**, 1678–1681.

Dauter, Z. & Adamiak, D. A. (2001). *Acta Cryst.* D**57**, 990–995.

Dauter, Z., Dauter, M., de La Fortelle, E., Bricogne, G. & Sheldrick, G. M. (1999). *J. Mol. Biol.* **289**, 83–92.

Dauter, Z., Dauter, M. & Dodson, E. J. (2002). *Acta Cryst.* D**58**, 494–506.

Debreczeni, J. É., Bunkóczi, G., Girmann, B. & Sheldrick, G. M. (2003). *Acta Cryst.* D**59**, 393–395.

Debreczeni, J. É., Girmann, B., Zeeck, A., Krätzner, R. & Sheldrick, G. M. (2003). *Acta Cryst.* D**59**, 2125–2132.

Devesse, L., Smirnova, I., Lönneborg, R., Kapp, U., Brzezinski, P., Leonard, G. A. & Dian, C. (2011). *Mol. Microbiol.* **81**, 354–367.

Doan, D. N. P. & Dokland, T. (2003). *Structure*, **11**, 1445–1451.

Einspahr, H. M. & Weiss, M. S. (2011). *International Tables for Crystallography*, Vol. *F*, 2nd online ed., edited by E. Arnold, D. M. Himmel & M. G. Rossmann, pp. 64–74. Chester: International Union of Crystallography.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.

Enroth, C. & Ake, S. (2008). *Biochim. Biophys. Acta*, **1784**, 379–384.

Evans, P. R. (2011). *Acta Cryst.* D**67**, 282–292.

Evans, P. R. & Murshudov, G. N. (2013). *Acta Cryst.* D**69**, 1204–1214.

Fabian, M. R., Frank, F., Rouya, C., Siddiqui, N., Lai, W. S., Karetnikov, A., Blackshear, P. J., Nagar, B. & Sonenberg, N. (2013). *Nature Struct. Mol. Biol.* **20**, 735–739.

Fukakusa, S., Kawahara, K., Nakamura, S., Iwashita, T., Baba, S., Nishimura, M., Kobayashi, Y., Honda, T., Iida, T., Taniguchi, T. & Ohkubo, T. (2012). *Acta Cryst.* D**68**, 1418–1429.

Gordon, E. J., Leonard, G. A., McSweeney, S. & Zagalsky, P. F. (2001). *Acta Cryst.* D**57**, 1230–1237.

Goulet, A., Vestergaard, G., Felisberto-Rodrigues, C., Campanacci, V., Garrett, R. A., Cambillau, C. & Ortiz-Lombardía, M. (2010). *Acta Cryst.* D**66**, 304–308.

Guimarães, B. G., Sanfelici, L., Neuenschwander, R. T., Rodrigues, F., Grizolli, W. C., Raulik, M. A., Piton, J. R., Meyer, B. C., Nascimento, A. S. & Polikarpov, I. (2009). *J. Synchrotron Rad.* **16**, 69–75.

Hendrickson, W. A. & Teeter, M. M. (1981). *Nature (London)*, **290**, 107–113.

Huet, J., Teinkela Mbosso, E. J., Soror, S., Meyer, F., Looze, Y., Wintjens, R. & Wohlkönig, A. (2013). *Acta Cryst.* D**69**, 2017–2026.

Ivanov, I., Crepin, T., Jamin, M. & Ruigrok, R. W. H. (2010). *J. Virol.* **84**, 3707–3710.

Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.

Klock, H. E. *et al.* (2005). *Proteins*, **61**, 1132–1136.

Lakomek, K., Dickmanns, A., Mueller, U., Kollmann, K., Deuschl, F., Berndt, A., Lübke, T. & Ficner, R. (2009). *Acta Cryst.* D**65**, 220–228.

Lartigue, A., Gruez, A., Briand, L., Blon, F., Bézirard, V., Walsh, M., Pernollet, J. C., Tegoni, M. & Cambillau, C. (2004). *J. Biol. Chem.* **279**, 4459–4464.

Li, D. *et al.* (2014). *Cryst. Growth Des.* **14**, 2034–2047.

Li, S. *et al.* (2002). *J. Biol. Chem.* **277**, 48596–48601.

Li, Y. *et al.* (2008). *Proteins*, **4**, 2109–2113.

Liu, K., Guo, Y., Liu, H., Bian, C., Lam, R., Liu, Y., Mackenzie, F., Rojas, L. A., Reinberg, D., Bedford, M. T., Xu, R. M. & Min, J. (2012). *PLoS One*, **7**, e30375.

Liu, Q., Dahmane, T., Zhang, Z., Assur, Z., Brasch, J., Shapiro, L., Mancia, F. & Hendrickson, W. A. (2012). *Science*, **336**, 1033–1037.

Liu, Q., Guo, Y., Chang, Y., Cai, Z., Assur, Z., Mancia, F., Greene, M. I. & Hendrickson, W. A. (2014). *Acta Cryst.* D**70**, 2544–2557.

Liu, Z.-J., Vysotski, E. S., Vysotski, E. S., Chen, C.-J., Rose, J. P., Lee, J. & Wang, B.-C. (2000). *Protein Sci.* **9**, 2085–2093.

Machius, M., Brautigam, C. A., Tomchick, D. R., Ward, P., Otwinowski, Z., Blevins, J. S., Deka, R. K. & Norgard, M. V. (2007). *J. Mol. Biol.* **373**, 681–694.

Micossi, E., Hunter, W. N. & Leonard, G. A. (2002). *Acta Cryst.* D**58**, 21–28.

Moynie, L., Schnell, R., McMahon, S. A., Sandalova, T., Boulkerou, W. A., Schmidberger, J. W., Alphey, M., Cukier, C., Duthie, F., Kopec, J., Liu, H., Jacewicz, A., Hunter, W. N., Naismith, J. H. & Schneider, G. (2013). *Acta Cryst.* F**69**, 25–34.

Mueller, U., Darowski, N., Fuchs, M. R., Förster, R., Hellmig, M., Paithankar, K. S., Pühringer, S., Steffien, M., Zocher, G. & Weiss, M. S. (2012). *J. Synchrotron Rad.* **19**, 442–449.

Mueller, J. W., Link, N. M., Matena, A., Hoppstock, L., Rüppel, A., Bayer, P. & Blankenfeldt, W. (2011). *J. Am. Chem. Soc.* **133**, 20096–20099.

Mueller-Dieckmann, C., Panjikar, S., Schmidt, A., Mueller, S., Kuper, J., Geerlof, A., Wilmanns, M., Singh, R. K., Tucker, P. A. & Weiss, M. S. (2007). *Acta Cryst.* D**63**, 366–380.

Mueller-Dieckmann, C., Panjikar, S., Tucker, P. A. & Weiss, M. S. (2005). *Acta Cryst.* D**61**, 1263–1272.

Müller, J. J., Weiss, M. S. & Heinemann, U. (2011). *Acta Cryst.* D**67**, 936–944.

Ni, S., Benning, M. M., Smola, M. J., Feldmann, E. A. & Kennedy, M. A. (2008). *Proteins*, **74**, 794–798.

Ni, S., Sheldrick, G. M., Benning, M. M. & Kennedy, M. A. (2008). *J. Struct. Biol.* **165**, 47–52.

Okada, U., Kondo, K., Hayashi, T., Watanabe, N., Yao, M., Tamura, T. & Tanaka, I. (2008). *Acta Cryst.* D**64**, 198–205.

Olsen, J. G., Flensburg, C., Olsen, O., Bricogne, G. & Henriksen, A. (2004). *Acta Cryst.* D**60**, 250–255.

Pal, A., Debreczeni, J. É., Sevvana, M., Gruene, T., Kahle, B., Zeeck, A. & Sheldrick, G. M. (2008). *Acta Cryst.* D**64**, 985–992.

Phillips, J. D., Whitby, F. G., Warby, C. A., Labbe, P., Yang, C., Pflugrath, J. W., Ferrara, J. D., Robinson, H., Kushner, J. P. & Hill, C. P. (2004). *J. Biol. Chem.* **279**, 38960–38968.

Pieren, M., Prota, A. E., Ruch, C., Kostrewa, D., Wagner, A., Biedermann, K., Winkler, F. K. & Ballmer-Hofer, K. (2006). *J. Biol. Chem.* **281**, 19578–19587.

Popovic, K., Holyoake, J., Pomès, R. & Privé, G. G. (2012). *Proc. Natl Acad. Sci. USA*, **109**, 2908–2912.

Rajasekaran, D., Fan, C., Meng, W., Pflugrath, J. W. & Lolis, E. J. (2013). *Proteins*, **82**, 708–716.

Rodamilans, B., Muñoz, I. G., Bragado-Nilsson, E., Sarrias, M. R., Padilla, O., Blanco, F. J., Lozano, F. & Montoya, G. (2007). *J. Biol. Chem.* **282**, 12669–12677.

Ru, H., Zhao, L., Ding, W., Jiao, L., Shaw, N., Liang, W., Zhang, L., Hung, L.-W., Matsugaki, N., Wakatsuki, S. & Liu, Z.-J. (2012). *Acta Cryst.* D**68**, 521–530.
Rudiño-Piñera, E., Ravelli, R. B. G., Sheldrick, G. M., Nanao, M. H., Korostelev, V. V., Werner, J. M., Schwarz-Linek, U., Potts, J. R. & Garman, E. F. (2007). *J. Mol. Biol.* **368**, 833–844.
Sarma, G. N. & Karplus, P. A. (2006). *Acta Cryst.* D**62**, 707–716.
Sarma, G. N., Kinderman, F. S., Kim, C., von Daake, S., Chen, L., Wang, B.-C. & Taylor, S. S. (2010). *Structure*, **18**, 155–166.
Sharma, P., Ishiyama, N., Nair, U., Li, W., Dong, A., Miyake, T., Wilson, A., Ryan, T., MacLennan, D. H., Kislinger, T., Ikura, M., Dhe-Paganon, S. & Gramolini, A. O. (2012). *FEBS J.* **279**, 3952–3964.
Shinobu, A., Palm, G. J., Schierbeek, A. J. & Agmon, N. (2010). *J. Am. Chem. Soc.* **132**, 11093–11102.
Stegmann, C. M., Seeliger, D., Sheldrick, G. M., de Groot, B. L. & Wahl, M. C. (2009). *Angew. Chem. Int. Ed. Engl.* **48**, 5207–5210.
Terwilliger, T. C., Adams, P. D., Read, R. J., McCoy, A. J., Moriarty, N. W., Grosse-Kunstleve, R. W., Afonine, P. V., Zwart, P. H. & Hung, L.-W. (2009). *Acta Cryst.* D**65**, 582–601.
Terwilliger, T. C., Grosse-Kunstleve, R. W., Afonine, P. V., Moriarty, N. W., Zwart, P. H., Hung, L.-W., Read, R. J. & Adams, P. D. (2008). *Acta Cryst.* D**64**, 61–69.
Thomas, S. R., McTamney, P. M., Adler, J. M., LaRonde-LeBlanc, N. & Rokita, S. E. (2009). *J. Biol. Chem.* **284**, 19659–19667.
Thorsell, A. G., Persson, C., Gräslund, S., Hammarström, M., Busam, R. D. & Hallberg, B. M. (2009). *Proteins*, **77**, 242–246.
Trempe, J.-F., Brown, N. R., Lowe, E. D., Gordon, C., Campbell, I. D., Noble, M. E. M. & Endicott, J. A. (2005). *EMBO J.* **24**, 3178–3189.
Vasur, J., Kawai, R., Larsson, A. M., Igarashi, K., Sandgren, M., Samejima, M. & Ståhlberg, J. (2006). *Acta Cryst.* D**62**, 1422–1429.
Wang, J., Dauter, M. & Dauter, Z. (2006). *Acta Cryst.* D**62**, 1475–1483.
Watanabe, N., Kitago, Y., Tanaka, I., Wang, J., Gu, Y., Zheng, C. & Fan, H. (2005). *Acta Cryst.* D**61**, 1533–1540.
Weiss, M. S. (2001). *J. Appl. Cryst.* **34**, 130–135.
Weiss, M. S., Mander, G., Hedderich, R., Diederichs, K., Ermler, U. & Warkentin, E. (2004). *Acta Cryst.* D**60**, 686–695.
Yogavel, M., Tripathi, T., Gupta, A., Banday, M. M., Rahlfs, S., Becker, K., Belrhali, H. & Sharma, A. (2014). *Acta Cryst.* D**70**, 91–100.
Zebisch, M., Schäfer, P., Lauble, P. & Sträter, N. (2013). *Acta Cryst.* F**69**, 257–262.
Zhu, J.-Y., Fu, Z.-Q., Chen, L., Xu, H., Chrzas, J., Rose, J. & Wang, B.-C. (2012). *Acta Cryst.* D**68**, 1242–1252.